The power of symmetric extensions for entanglement detection

Miguel Navascués, [1] Masaki Owari,[1,2] Martin B. Plenio,[1,2]
[1] *Institute for Mathematical Sciences, 53 Prince's Gate, Imperial College London, London SW7 2PG, UK*
[2] *QOLS, Blackett Laboratory, Imperial College London, London SW7 2BW, UK*

In this paper, we present new progress on the study of the symmetric extension criterion for separability. First, we show that a perturbation of order $O(1/N)$ is sufficient and, in general, necessary to destroy the entanglement of any state admitting an N Bose symmetric extension. On the other hand, the minimum amount of local noise necessary to induce separability on states arising from N Bose symmetric extensions with Positive Partial Transpose (PPT) decreases at least as fast as $O(1/N^2)$. From these results, we derive upper bounds on the time and space complexity of the weak membership problem of separability when attacked via algorithms that search for PPT symmetric extensions. Finally, we show how to estimate the error we incur when we approximate the set of separable states by the set of (PPT) N-extendable quantum states in order to compute the maximum average fidelity in pure state estimation problems, the maximal output purity of quantum channels, and the geometric measure of entanglement.

PACS numbers:

I. INTRODUCTION

The separability problem, that is, the problem to determine whether a given quantum state is separable or entangled, is one of the most fundamental problems in Entanglement Theory [1]. Starting from the famous PPT (Positive Partial Transpose) criterion [2], nowadays we have an enormous number of different separability criteria to choose from (see the citation lists of review papers in this topic [1, 3, 4, 5, 6, 7, 8]). Among all known separability criteria, those based on "symmetric extensions" and "PPT symmetric extensions" (i.e., symmetric extensions with an additional PPT constraint), as conceived by Doherty et al. [9, 10], are considered to be among the most powerful [6]. These criteria rely on the fact that any set of N-symmetrically extendable states (PPT or not) converges to a set of separable states in the limit of $N \to \infty$, as first noticed by Raggio and Werner [11, 12], although it also follows from the Quantum de Finetti theorem [13]. Since both the set of N-symmetrically extendable states and the set of N-PPT symmetrically extendable states can be characterized by Semidefinite Programming [14], a well-known optimization problem for which many free solvers are available (like the MATLAB toolbox SeDuMi[15]), these tests are not only powerful, but also easy to implement. This explains why, over all known numerical methods, the algorithms created by Doherty, Parrilo and Spedalieri (DPS) are the most popular in the Quantum Information community (notice, however, that there exist other methods for entanglement detection based on Semidefinite Programming besides the DPS criterion [16, 17]).

This family of schemes has, though, an important drawback: in this approach, in order to conclude that a given state ρ is entangled, it is enough to find an N such that ρ does not belong to the set of N-(PPT) symmetric extendable states. On the other hand, in order to show that a given state is separable, we would have to prove that it admits an N-(PPT) symmetric extension *for all natural numbers* N. The DPS method then becomes useless: since we always operate under finite time and memory constraints, all we can do in practice is to check for the existence of N-(PPT) symmetric extensions for N less or equal than some finite number N_0. If the state ρ under analysis happened to admit an N_0 (PPT) symmetric extension, we could thus not conclude anything about its separability.

Hulke and Bruss [18] tried to solve the issue by providing a complementary criterion designed to detect separability instead of entanglement, to be implemented at the same time as the DPS criterion. Unfortunately, the time complexity of that other method scales superexponentially with the dimension of the subsystems involved [6]. The reduced speed of convergence of the resulting two-way algorithm (much smaller than that of the DPS criterion) thus makes it unsuitable to study quantum correlations in high dimensional systems.

Besides, there is a more elegant way to approach the problem.

In a recent work, Ioannou observed that, even if a state happens to have an N_0-(PPT) symmetric extension, we can at least bound the distance between such state and the set of separable states in terms of N_0 [6]. In the language of Computer Science, this means that the "truncated" DPS criterion allows to solve an instance of an approximate separability problem, the weak membership problem of separability ($WMEM(\bar{S})$). Ioannou therefore provided an upper bound on the full time-complexity of the algorithm for WMEM based on symmetric extension criteria.

But even after Ioannou's work, an open question remains to be solved. The PPT symmetric extension criterion is considered to be stronger than the symmetric extension criterion [9, 10]. By definition, it is actually at least as strong as the symmetric extension criterion in the sense that a N-PPT symmetrically extendable state

is N-symmetrically extendable. However, so far, there are no results that quantify *how strong* the additional PPT constraint makes the DPS criterion. In particular, since the additional PPT constraint increases quadratically the size of the matrices that define the Semidefinite Programming problem, there still remains the possibility that the PPT criterion just makes the DPS algorithm slower for *WMEM*($\bar{S}$). In order to make this point clear, a similar analysis as Ioannou's should be done for the PPT-symmetric extension criteria. Since Ioannou's analysis is based on the finite quantum de Finetti theorem [19, 20] and there exists no similar theorem for states satisfying the PPT constraint, there is no straightforward extension of Ioannou's work to the PPT symmetric extension criterion.

In this paper, by analyzing these criteria in more detail, we extend Ioannou's result to account for the PPT condition.

The structure of this article is as follows: in Section II we will give the reader a detailed explanation of the DPS criterion and introduce the basic notation that will be used in the paper. Then we will move on to present the main result of this article, namely, an upper bound on the amount of noise needed to make the DPS states separable. This will allow us to compute upper bounds on the entanglement robustness of these states, and on their distance to the set of separable states. We will also briefly discuss how close our bounds are to being optimal. In Section IV, we will use the previous results to analyze the computational complexity of solving the weak membership problem of separability through the DPS criterion. In particular, we will show that the PPT constraint in the DPS criterion reduces the dominant factor of the upper bound on the time complexity from $(k_1/\delta)^{6d_B}$ to $(k_2/\delta)^{4d_B}$, where δ is the accuracy parameter of *WMEM*($\bar{S}$). In Section V we will bound the speed of convergence of the DPS criterion when applied to compute the optimal fidelity in state estimation problems, the output purity of quantum channels and the geometric entanglement of arbitrary states. There we will perform some numerical tests to have a grasp at the actual speed of convergence of the DPS criterion, as opposed to our analytical upper bounds on it. In Sections VI, VII we will give the proof of the main theorem and explain how it can be extended to deal with the multipartite case. Afterwards, we will also show a very simple method to bound the entanglement of general PPT states. Finally, Section IX will present our conclusions.

II. THE DPS CRITERION

The Doherty-Parrilo-Spedalieri (DPS) criterion for entanglement detection [9] is a numerical algorithm that, combining the aforementioned results [11, 12, 13] on N-extendibility with convex optimization methods, allows to characterize the set S of separable operators up to arbitrary precision. The criterion arises from the following observation: if $\Lambda_{AB} \in S$, then, by definition, it belongs to the cone of bipartite product states, i.e.,

$$\Lambda_{AB} = \sum_i p_i |u_i\rangle\langle u_i| \otimes |v_i\rangle\langle v_i|, \tag{1}$$

with $p_i \geq 0$ for all i.

Once this decomposition is known, we can define a uniparametric family of operators $\Lambda_{AB^N} \in B(\mathcal{H}_A \otimes \mathcal{H}_B^{\otimes N})$ by tensoring N times the last part:

$$\Lambda_{AB^N} \equiv \sum_i p_i |u_i\rangle\langle u_i| \otimes |v_i\rangle\langle v_i|^{\otimes N}. \tag{2}$$

Let us study the properties of the newly defined operators: first of all, from the above definition it is clear that they are all positive semidefinite. Also, from (2) it can be seen that tracing out the last $N-1$ systems we recover the initial operator, i.e., $\mathrm{tr}_{B^{N-1}}(\Lambda_{AB^N}) = \Lambda_{AB}$, and that the last N systems are invariant under the action of the permutation group. Finally, when viewed as an $N+1$-partite system, Λ_{AB^N} is multiseparable, and therefore must remain positive semidefinite under the partial transposition of any bipartition of these systems.

For simplicity, we will incorporate all these properties in a single definition:

Definition 1. *Bose symmetric extensions (BSE) Let $\Lambda_{AB} \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be a non-negative operator. We will say that $\Lambda_{AB^N} \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B^{\otimes N})$ is an N Bose symmetric extension (BSE) of Λ_{AB} iff:*

1. $\Lambda_{AB^N} \geq 0$.
2. $tr_{B^{N-1}}(\Lambda_{AB^N}) = \Lambda_{AB}$.
3. *Λ_{AB^N} is Bose symmetric, i.e., $\Lambda_{AB^N}(\mathbb{I}_A \otimes P^N_{sym}) = \Lambda_{AB^N}$, where P^N_{sym} denotes the symmetric projector of N particles.*

In case Λ_{AB^N} is PPT with respect to all or some of its bipartitions $AB^K|B^{N-K}$, we will call it a *PPT Bose symmetric extension* (PPT BSE) of Λ_{AB}.

From what we have seen, it is clear that, if Λ_{AB} is a separable operator, then there exists an N (PPT) BSE of Λ_{AB} for any N. Since (PPT) Bose symmetric extensions are defined through linear matrix inequalities, the problem of determining whether a given state Λ_{AB} admits one or not can be cast as a semidefinite program (SDP) [14], and therefore can be solved efficiently for fixed N and varying dimensions. The DPS criterion consists precisely in, given an operator Λ_{AB} whose separability is at stake, check for the existence of N (PPT) Bose symmetric extensions for different values of N.

A hierarchy of separability tests arises then naturally: if some operator Λ_{AB} does not admit a (PPT) Bose symmetric extension for some N (i.e., it does not pass the N^{th} test), then it has to be entangled. If, on the contrary, such extension exists, then we would go for the

$(N+1)^{th}$ test, that is, we would search for $N+1$ (PPT) Bose symmetric extensions of Λ_{AB}. This last test would be in general more restrictive than the previous one, since for any $N+1$ (PPT) Bose symmetric extension $\Lambda_{AB^{N+1}}$ of Λ_{AB} we can obtain an N (PPT) Bose symmetric extension by tracing out the last system.

Doherty et al. [10] showed that the previous hierarchy completely characterizes the set of separable operators, in the sense that for any entangled positive operator Λ_{AB} there exists an N such that Λ_{AB} does not admit an N Bose symmetric extension.

We will now introduce a notation that will be used for the rest of the article: S^N will denote the *cone* of all bipartite operators that have an N BSE, and S_p^N will be understood as the set of all unnormalized quantum states that admit an N BSE that is *PPT with respect to the bipartition* $AB^{\lceil N/2\rceil}|B^{\lfloor N/2\rfloor}$. In case we also demand normalization, we will be dealing with the sets of states $\bar{S}^N, \bar{S}_p^N$. The elements of the previous four sets will be called N-(PPT) symmetrically extendable operators, or states, if normalized, or just DPS operators or states. Our previous discussion can then be summarized as

$$\begin{aligned} &S^1 \supset S^2 \supset S^3 \supset ... \supset S, \\ &S_p^1 \supset S_p^2 \supset S_p^3 \supset ... \supset S, \\ &\lim_{N\to\infty} S^N, S_p^N = S. \end{aligned} \tag{3}$$

Note that $S^1 = S_p^1$ ($\bar{S}^1 = \bar{S}_p^1$) is the set of all positive semidefinite operators (states).

Before ending this section, we would like to point out one additional fact. As we already explained in the introduction, when we use the DPS criterion in practice, it is not possible to conclude with certainty that a given state is separable. However, in the PPT case, by checking some rank constraints on the density matrices output by the computer, we can *sometimes* conclude separability in a finite number of steps. In that case, we will say that the PPT BSE presents a *rank loop*. We will make use of rank loops in Section V in order to estimate the accuracy of our upper bounds on the error we introduce when we perform linear optimizations over the sets S^N or S_p^N instead of S in state estimation problems. A detailed explanation of this criterion for optimality can be found in Appendix B.

III. CHARACTERIZATION OF S^N AND S_p^N

We have seen that the sequences of sets (S^N), (S_p^N) tend to the set S in the limit $N \to \infty$. Intuitively, this means that, for $N >> 1$, any state ρ_{AB} belonging to one of these sets must be either separable, or, at least, very close to a separable state. It seems thus plausible that the little entanglement such states may possess could be destroyed by some very attenuated local noise. One of the most simple noise models one can think of is depolarization, where a quantum state is turned into white noise with probability p. The action of the depolarizing channel $\Omega^{(p)}$ over some state $\rho \in B(\mathcal{H})$ is given by

$$\Omega^{(p)}(\rho) = (1-p)\rho + p\frac{\mathbb{I}}{d}, \tag{4}$$

where d is the dimension of the Hilbert space $\mathcal{H}$. Given any bipartite quantum state ρ_{AB}, shared by Alice and Bob, we could thus define its *critical disentangling probability* $p_c(\rho_{AB})$ as the minimum probability with which one of the parties, say Bob, would have to prepare the maximally mixed state in his subsystem in order to disentangle it from Alice's. That is,

$$p_c(\rho_{AB}) = \min\{p : \mathbb{I}_A \otimes \Omega_B^{(p)}(\rho_{AB}) \in \bar{S}\}. \tag{5}$$

Similarly, we can define the critical disentangling probability of a set of states W as the maximum of all $p_c(\rho)$ for all $\rho \in W$. Clearly, $p_c \leq 1$ for all states, although this bound can be greatly improved if the dimensionality of Bob's system is small, as we shall see.

In this section, we will give upper bounds on this critical probability valid for any state in $\bar{S}^N$ (or $\bar{S}_p^N$). Then, by means of these results, we will provide several upper bounds on the speed of convergence of $\bar{S}^N$ and $\bar{S}_p^N$ to $\bar{S}$.

Before proceeding, though, a remark on notation: in this article, we will be mainly concerned with linear operators or quantum states acting over a bipartite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, and all the formulas and bounds that we will derive in this section and the following three will involve the dimension of the Hilbert space $\mathcal{H}_B$ where the symmetric extensions are to be made. For the sake of clarity, we will therefore introduce the notation $d \stackrel{\text{def}}{=} \dim \mathcal{H}_B$.

The following theorems will play a key role in deriving most of the results of this paper.

Theorem 2.

$$p_c(\bar{S}^N) \leq \frac{d}{N+d}. \tag{6}$$

In other words: for any operator $\Lambda_{AB} \in S^N$, the positive semidefinite operator

$$\tilde{\Lambda}_{AB} \equiv \frac{N}{N+d}\Lambda_{AB} + \frac{1}{N+d}\Lambda_A \otimes \mathbb{I}_B \tag{7}$$

is separable.

Theorem 3. *Define g_N (or $g_N^{(d)}$ in case d is ambiguous) as*

$$\begin{aligned} g_N = \quad & \min\{1-x : P_{N/2+1}^{(d-2,0)}(x) = 0\} \text{ for } N \text{ even}, \\ & \min\{1-x : P_{(N+1)/2}^{(d-2,1)}(x) = 0\} \text{ for } N \text{ odd}, \end{aligned} \tag{8}$$

with $P_n^{(\alpha,\beta)}(x)$ being the Jacobi Polynomials [21].

Then,

$$p_c(\bar{S}_p^N) \leq \frac{d}{2(d-1)} g_N. \quad (9)$$

That is, for any $\Lambda_{AB} \in S_p^N$, the positive semidefinite operator

$$\tilde{\Lambda}_{AB} \equiv (1 - \frac{d}{2(d-1)} g_N)\Lambda_{AB} + \frac{1}{2(d-1)} g_N \Lambda_A \otimes \mathbb{I}_B \quad (10)$$

is separable.

The proof of these two theorems is given in Section VI, where a separable decomposition for the states (7), (10) is also provided. Also, it is worth mentioning that, in both cases, $\tilde{\Lambda}_A = \Lambda_A$.

Notice that, in Theorem 3, g_N is defined in terms of the greatest root of Jacobi polynomials. The properties of the roots of Jacobi polynomials have been studied for quite time [21]. This allows us to derive an expression for the asymptotic behavior of g_N:

$$g_N \approx 2\left(\frac{j_{d-2,1}}{N}\right)^2, \text{ for } N >> 1, \quad (11)$$

$$\approx 2\left(\frac{d + 1.856 d^{1/3} + O(d^{-1/3})}{N}\right)^2, \quad \text{for } N \gg d \gg 1, \quad (12)$$

where $j_{n,1}$ is the first positive zero of the Bessel function $J_n(y)$.

How far can then the states in $\bar{S}^N, \bar{S}_p^N$ be from the set $\bar{S}$ of separable states? A way to answer this question could be to bound the maximum possible entanglement of such states.

The *robustness of entanglement* of a state ρ is defined as the minimum amount of separable noise needed to destroy the entanglement of such a state [22]:

$$R(\rho) \stackrel{\text{def}}{=} \min_{\lambda} \{\lambda : \exists \sigma \in \bar{S}, \text{ s.t. } \frac{\rho + \lambda\sigma}{1+\lambda} \in S\}. \quad (13)$$

The robustness of entanglement is also an upper bound on the *global robustness of entanglement* $R_G(\rho)$ [22], defined by allowing σ to be an arbitrary normalized quantum state in the above expression. And the global robustness of entanglement is, in turn, lower bounded by several other entanglement measures, like the negativity, the geometric measure of entanglement and the relative entropy of entanglement [22, 23, 24, 25, 26]. Any non trivial upper bound on the entanglement robustness of the states in $\bar{S}^N$ and $\bar{S}_p^N$ could thus retrieve a lot of information.

The following corollary follows straightforwardly from theorems 2 and 3.

Corollary 4. *Any $\rho \in \bar{S}^N$ satisfies*

$$R(\rho) \leq \frac{d-1}{N}. \quad (14)$$

Similarly, any $\rho \in \bar{S}_p^N$ satisfies

$$R(\rho) \leq \frac{g_N}{2 - \frac{d}{d-1} g_N} \approx \left(\frac{d}{N}\right)^2. \quad (15)$$

To see why, suppose that ρ is normalized and use formulas (7), (10) to express $\tilde{\rho}$ (i.e., $\tilde{\Lambda}_{AB}$) in each case as a convex sum of the non negative operators ρ and $\sigma \equiv \frac{1}{d-1}(\rho_A \otimes \mathbb{I}_B - \tilde{\rho})$. Then, notice that, since $\tilde{\rho}_A = \rho_A$ and $\tilde{\rho}$ is separable, then σ must also be a separable operator [61]

Theorems 2 and 3 also allow to obtain bounds on the distance between the states in $\rho_{AB} \in S^N, S_p^N$ and the set of separable states $\bar{S}$.

Corollary 5. *For any $\rho \in \bar{S}^N$, there exist $\tilde{\rho} \in \bar{S}$ such that*

$$\|\rho - \tilde{\rho}\|_1 \leq \frac{2(d-1)}{N+d-1}, \quad (16)$$

$$\|\rho - \tilde{\rho}\|_\infty \leq \frac{d-1}{N+d-1}, \quad (17)$$

$$\|\rho - \tilde{\rho}\|_F = \frac{d}{N+d}\sqrt{tr(\rho^2) - \frac{tr(\rho_A^2)}{d}}, \quad (18)$$

where $\|\cdot\|_1$, $\|\cdot\|_\infty$ and $\|\cdot\|_F$ are the trace, the operator and the Frobenius norm, respectively.

Similarly, for any $\rho \in \bar{S}_p^N$ (and $N \geq 2$), there exists a state $\tilde{\rho} \in \bar{S}$ such that

$$\|\rho - \tilde{\rho}\|_1 \leq g_N, \quad (19)$$

$$\|\rho - \tilde{\rho}\|_\infty \leq g_N/2, \quad (20)$$

$$\|\rho - \tilde{\rho}\|_F = \frac{d g_N}{2d-2}\sqrt{tr(\rho^2) - \frac{tr(\rho_A^2)}{d}}. \quad (21)$$

Proof. Here we give the proof for the bounds on the trace and operator norm. The proof for the Frobenius norm is omitted, since it is similar and simpler.

Let $\rho \in \bar{S}^N$. Them Theorem 2 implies that there exists $\tilde{\rho} \in \bar{S}$, with $\tilde{\rho}_A = \rho_A$, such that:

$$\rho - \tilde{\rho} = \frac{d-1}{N+d-1}\rho - \frac{1}{N+d-1}(\rho_A \otimes \mathbb{I}_B - \tilde{\rho}). \quad (22)$$

Using the triangle inequality, we have that

$$\|\rho - \tilde{\rho}\|_1 \leq \frac{d-1}{N+d-1}\|\rho\|_1 + \\ + \frac{1}{N+d-1}\|(\rho_A \otimes \mathbb{I}_B - \tilde{\rho})\|_1 = \frac{2(d-1)}{N+d-1}, \quad (23)$$

where in the last step we used once more the fact that $\rho_A \otimes \mathbb{I}_B - \tilde{\rho}$ is separable (and, therefore, positive). Relation (16) is thus proven.

For the operator norm, let $u_+(u_-)$ be the eigenvector corresponding to the maximum (minimum) eigenvalue of $\rho-\tilde{\rho}$. It follows that

$$\|\rho-\tilde{\rho}\|_\infty = \max(\mathrm{tr}\{(\rho-\tilde{\rho})|u_+\rangle\langle u_+|\}, \mathrm{tr}\{(\tilde{\rho}-\rho)|u_-\rangle\langle u_-|\}). \tag{24}$$

On the other hand,

$$\mathrm{tr}\{(\rho-\tilde{\rho})|u_+\rangle\langle u_+|\} = \tfrac{d-1}{N+d-1}\mathrm{tr}\{\rho|u_+\rangle\langle u_+|\} - \\ -\tfrac{1}{N+d-1}\mathrm{tr}\{(\rho_A\otimes\mathbb{I}_B-\tilde{\rho})|u_+\rangle\langle u_+|\} \leq \tfrac{d-1}{N+d-1}, \tag{25}$$

and

$$\mathrm{tr}\{(\tilde{\rho}-\rho)|u_-\rangle\langle u_-|\} = -\tfrac{d-1}{N+d-1}\mathrm{tr}\{\rho|u_-\rangle\langle u_-|\} + \\ +\tfrac{1}{N+d-1}\mathrm{tr}\{(\rho_A\otimes\mathbb{I}_B-\tilde{\rho})|u_-\rangle\langle u_-|\} \leq \tfrac{d-1}{N+d-1}. \tag{26}$$

The first part of the corollary has been proven.

If $\rho\in\bar{S}_p^N$ and $N\geq 2$, then ρ can be seen to be PPT. Since the PPT criterion implies the reduction criterion [27, 28], we have that $\rho_A\otimes\mathbb{I}_B-\rho\geq 0$. This observation, combined with the techniques used to derive the first set of relations, allows to prove the second one. □

The above corollaries can be reformulated as:

Corollary 6. *Suppose $\bar{S}(\delta)$ is a δ-neighbor of the set of all separable states $\bar{S}$ in terms of the trace distance:*

$$\bar{S}(\delta) \stackrel{\mathrm{def}}{=} \bigcup_{\rho\in\bar{S}} \left\{\sigma\in\bar{S}^1 \mid \|\rho-\sigma\|\leq\delta\right\} \tag{27}$$

(remember that $\bar{S}^1$ is the set of all quantum states in $\mathcal{H}_A\otimes\mathcal{H}_B$).
Then, the following relations hold:

$$\bar{S}^N \subset \bar{S}\left(\frac{2(d-1)}{N+d-1}\right)\approx\bar{S}\left(2\frac{d}{N}\right), \tag{28}$$

$$\bar{S}_p^N \subset \bar{S}(g_N)\approx\bar{S}\left(2\left(\frac{d}{N}\right)^2\right), \tag{29}$$

where the approximations are granted to hold in the limit $N\gg d\gg 1$.

This corollary suggests that the upper bounds for $\bar{S}_p^N$ converge quadratically faster than those for $\bar{S}^N$. In other words, if these bounds were optimum, then we would have proven that the additional PPT constrain gives the DPS criterion a quadratic speed-up.

It is then natural to wonder if such bounds are indeed optimal. We will argue that at least the scaling of the upper bounds for $\bar{S}^N$ is correct, i.e., fixing d_A and d, the maximum possible entanglement robustness of any bipartite state ρ_{AB} arising from an N Bose symmetric extension scales with N as $O(1/N)$.

To see this, let $N=2K-1$, and consider the $N+1$ bipartite state given by

$$|\Psi_{AB^N}\rangle \equiv \frac{1}{C_K}\sum_{\mathrm{perm}} |\overbrace{0\cdot 0}^{K}\overbrace{1\cdot 1}^{K}\rangle, \tag{30}$$

where C_K is a normalization factor. Define now $\rho_{AB}\equiv \mathrm{tr}_{B^{N-1}}(|\Psi_{AB^N}\rangle\langle\Psi_{AB^N}|)$. Clearly, $\rho_{AB}\in S^N$. Now, it can be shown that

$$\rho_{AB} = \frac{K-1}{2(2K-1)}(|00\rangle\langle 00|+|11\rangle\langle 11|) + \\ +\frac{K}{2(2K-1)}(|01\rangle+|10\rangle)(\langle 01|+\langle 10|). \tag{31}$$

The partially transposed operator $\rho_{AB}^{T_B}$ has a negative eigenvalue $-1/2(2K-1)$ corresponding to the eigenvector $(|00\rangle-|11\rangle)/\sqrt{2}$, whose maximum Schmidt coefficient is $1/\sqrt{2}$. According to [22], this implies that $R(\rho_{AB})=1/(2K-1)=1/N$. The bound (14) is, therefore, tight for $d_A=d=2$. Since for any pair for Hilbert spaces $\mathcal{H}_A,\mathcal{H}_B$ of dimensions greater than 1 we can embed the previous family of states in $B(\mathcal{H}_A\otimes\mathcal{H}_B)$, it follows that the optimal upper bound on the entanglement robustness of partial traces of Bose symmetric extensions must scale as $O(1/N)$. On the other hand, the bound (15) guarantees that the corresponding value for $\bar{S}_p^N$ at least scales as $O(1/N^2)$, i.e., Theorem 3 allows to derive an upper bound for the entanglement robustness that decreases asymptotically faster than the optimal upper bound in the general Bose symmetric case.

Note that the above considerations also allow us to obtain a dimension-dependent lower bound on the maximum possible entanglement robustness R_{sup}^N of a state in $\bar{S}^N$. Following the lines of [38], consider the state $\sigma\equiv\rho_{AB}^{\otimes M}$, with ρ_{AB} given by equation (31). Clearly, $\sigma\in\bar{S}^N$, with $d_A=d_B=d=2^M$. As $-1/(2N)$ is the only negative eigenvalue of $\rho_{AB}^{T_B}$ and, therefore, the sum of its positive eigenvalues adds up to $1+1/(2N)$, the negativity of σ [39] (i.e., minus the sum of the negative eigenvalues of σ^{T_B}) can be seen equal to

$$\mathcal{N}(\sigma) = \sum_{j=0}^{\lfloor (M-1)/2\rfloor}\binom{M}{2j+1}\frac{\left(1+\frac{1}{2N}\right)^{M-2j-1}}{(2N)^{2j+1}} = \\ = \frac{[(1+\frac{1}{2N})+\frac{1}{2N}]^M-[(1+\frac{1}{2N})-\frac{1}{2N}]^M}{2} = \\ = \frac{(1+\frac{1}{N})^M-1}{2}\approx\frac{M}{2N}, \tag{32}$$

where the last approximation is valid in the limit of large N. Since $R(\sigma)\geq\mathcal{N}(\sigma)$ [23], it follows that $R(\sigma)\gtrapprox O(\log(d)/N)$. That is, for fixed dimension d, R_{sup}^N satisfies $O(\log(d)/N)\leq R_{\mathrm{sup}}^N\leq O(d/N)$.

IV. COMPUTATIONAL COMPLEXITY OF WMEM($\bar{S}$)

In this section, we will analyze the consequences of the previous results on separability from the point of view of Computer Science. Actually, there are several different ways to describe the separability problem as a computational problem [6]. We chose to focus our attention in an approximated separability problem called the *weak membership problem of separability*. This *"promise"* problem (as opposed to a *"decision"* problem) roughly consists on deciding the separability of a given state, but allowing an uncertainty parameterized by δ. In this Section we will derive upper bounds on the time and space complexity when we attack this problem via the DPS criterion.

The "*In-biased*" weak membership problem is defined as follows [6]:

Definition 7. *Weak membership problem of separability (WMEM($\bar{S}$))*
Given a bipartite quantum state $\rho \in \bar{S}^1$ and rational $\delta > 0$, assert either that

$$\rho \in \bar{S}(\delta) \text{ or} \tag{33}$$

$$\rho \notin \bar{S}, \tag{34}$$

where $\bar{S}(\delta)$ is a δ neighbor of $\bar{S}$, i.e., $\bar{S}(\delta) = \{\sigma \in \bar{S}^1 : \tilde{\sigma} \in \bar{S} \subset \bar{S}^1, \|\tilde{\sigma} - \sigma\|_1 \leq \delta\}$.

In the above definition, $\|\omega\|_1 = \mathrm{tr}(\sqrt{\omega\omega^\dagger})$, the trace norm of the operator ω, although, in principle, we could have chosen other norms or distance measures as an accuracy parameter.

WMEM(S) is, thus, an approximation of the conventional separability problem in the sense that an algorithm solving WMEM($\bar{S}$) may assert equation (33) for a state ρ_{AB} having just a small amount of entanglement. This approximated formalism is more practical than a non-approximated or exact formalism like EXACT-QSEP [6], because of the inevitable errors we incur in both numerical and experimental studies, that should somehow be accounted for in our analysis of separability. A fair amount of effort has been devoted to the study of the time complexity of WMEM($\bar{S}$), the most remarkable result being that, if $d_A \geq d_B$, then WMEM(S) is NP-hard whenever $1/\delta$ increases exponentially [40] or polynomially [41] with respect to d_B.

We will now proceed to evaluate the time complexity of WMEM($\bar{S}$) when solved through the DPS criterion. First, following the discussion of Doherty et al. [10], S^N can be characterized by a semidefinite program with $\left((\dim \mathcal{H}_{\mathrm{sym}}^N)^2 - d_B^2\right) d_A^2$ free variables and a matrix of size $(\dim \mathcal{H}_{\mathrm{sym}}^N) d_A$ on which we will impose the positivity constraint. On the other hand, for $\bar{S}_p^N$, the PPT constraint implies demanding positivity from an additional matrix of size $(\dim \mathcal{H}_{\mathrm{sym}}^{N/2})^2 d_A$. Since the time-complexity of an SDP with m variables and of matrix size n is $O(m^2 n^2)$ (with a small extra cost coming from an iteration of algorithms), the dominant factors for the asymptotic time-complexity of these tests can be written as

$$\text{Symmetric} : d_A^6 (\dim \mathcal{H}_{\mathrm{sym}}^{\overline{N_{\mathrm{sym}}}})^6 \tag{35}$$

$$\text{PPT symmetric} : d_A^6 (\dim \mathcal{H}_{\mathrm{sym}}^{\overline{N_{\mathrm{ppt}}}})^4 (\dim \mathcal{H}_{\mathrm{sym}}^{\overline{N_{\mathrm{ppt}}}/2})^4, \tag{36}$$

where $\overline{N_{\mathrm{sym}}}$ and $\overline{N_{\mathrm{ppt}}}$ are the sizes of the extensions needed to achieve a given accuracy parameter δ.

Thus, at this stage, even though $\bar{S}_p^N$ converges to $\bar{S}$ faster than $\bar{S}^N$, there still remains the possibility that the algorithm based on the sets $\{\bar{S}_p^N\}$ is slower than the one based on the sets $\{\bar{S}^N\}$, because of the increase in time complexity that arises from imposing positivity on the partially transposed operator. The following calculation will rule out this possibility.

From Eq. (28) of Corollary 6, we have that

$$\overline{N_{\mathrm{sym}}} \leqq \tfrac{(2-\delta)(d_B-1)}{\delta},$$
$$\overline{N_{\mathrm{ppt}}} \lessapprox \tfrac{\sqrt{2} j_{d_B-2,1}}{\sqrt{\delta}}. \tag{37}$$

Taking into account that $j_{d,1} \approx d + O(d^{1/3})$ [21], the final expressions for upper bounds of the time complexity with respect to one method and the other are

$$O\left(d_A^6 \left[\frac{2e}{\delta}\right]^{6d_B}\right), \quad \text{for } \bar{S}^N$$
$$O\left(d_A^6 \left[\frac{e^2}{\delta}\right]^{4d_B}\right), \quad \text{for } \bar{S}_p^N, \tag{38}$$

where we just wrote the dominant (exponential) terms and omitted all polynomially growing terms. Note that the scaling law derived for the non PPT DPS criterion is valid as long as the optimal bounds on the trace distance to the set of separable states scale as d_B/N. We conjecture that such is the case, although all our attempts to derive an analytical proof have failed so far. Under this assumption, the above formula thus shows that the criterion based on PPT BSEs indeed requires less steps than the one based on plain BSEs in order to solve WMEM($\bar{S}$) for a given accuracy δ.

The space complexity of both the plain DPS criterion and the PPT DPS criterion, though, is of the same type. This is because, although the PPT condition imposes (at least) a quadratic speedup in the speed of convergence, it also increases quadratically the size of the matrices involved in the SDP. Thus one effect cancels the other, and the size of the matrices needed in both cases to solve WMEM($\bar{S}$) up to a given precision δ is comparable for any value of d_B. It follows that, according to our bounds, in some situations it may be more convenient not to use the PPT condition in order to save memory space.

Our experience with the DPS method suggests, however, that this expectation is not realistic, but rather a consequence of the non optimality of the bounds implicit in Theorem 3. Actually, in practice, the algorithm based on PPT BSEs seems to have smaller space complexity than the one based on general BSEs.

A big underestimation of the role of the PPT condition in the DPS criterion could also explain why the bound (38) behaves much worse than the asymptotic expressions $(k/\delta)^{2d_B}$ derived in [6] for the performance of the algorithm conceived by Ioannou et al. for entanglement detection [42, 43]. Indeed, as we will see, our bounds on the distance between the sets $\bar{S}_p^N$ and the set of separable states are far from optimal, at least for small values of d_A. Therefore, a more refined analysis could in principle end up with a different scaling law for this distance, that would eventually lead to a much better estimate of the time complexity of methods based in PPT BSEs.

V. APPROXIMATE ALGORITHMS FOR STATE ESTIMATION, MAXIMUM OUTPUT PURITY, AND GEOMETRIC MEASURE OF ENTANGLEMENT

There are many relevant quantities in quantum information whose definition involves a linear optimization over a set of separable operators. The maximum average fidelity in state estimation problems [49, 50], the output purity of a quantum channel [29] or the geometric measure of entanglement [1] are examples of such quantities. In order to compute these functions, we could think of an approximate algorithm that optimized over the sets S^N or S_p^N instead of S, and it is easy to see that such an algorithm would give the correct answer in the limit of large N.

So far, we have seen how Theorems 2 and 3 can be used to derive bounds related to the separability problem. In this Section we will show how to use these same theorems to bound the precision of the approximate linear optimizations over the cone of separable operators mentioned above.

A. State Estimation Problems

In a *general* state estimation scenario, a source chooses with probability p_i a virtual quantum state Ψ_i that is encoded afterwards into another quantum state Ψ_i', to which we are given full access. The goal of the game is to measure our given state by means of a Positive Operator Valued Measure (POVM) $\{M_x\}_x$ and thus obtain a classical value x that we will use to make a guess ϕ_x on the original state Ψ_i, which from now on we will assume to be pure. In conventional estimation theory, we usually restrict the guess ϕ_x to be one of the original states $\{\Psi_i\}_i$ [49, 50]. In this section, however, we will consider the more general setting in which we are allowed to choose arbitrary states as a guess.

Being Ψ_i a pure state, the efficiency of the protocol as a whole can be parametrized in terms of the *average fidelity* f:

$$0 \leq f \equiv \sum_{i,x} p_i \mathrm{tr}(\Psi_i' M_x)\mathrm{tr}(\phi_x \Psi_i) \leq 1. \qquad (39)$$

And the state estimation problem consists on determining F, the maximum fidelity among all possible measure-and-prepare schemes (M_x, ϕ_x). Since F can be used as well to determine whether a given quantum channel can be simulated or not by an entanglement breaking channel, this problem is also referred to as the *Quantum benchmark problem* [51, 52, 53, 54, 55].

In [56], it is explained how to map the SE problem into a linear optimization over the set S of separable states, via the relation

$$F = \max\{\mathrm{tr}(\rho_{AB}\Lambda_{AB}) : \Lambda_{AB} \in S, \Lambda_A = \mathbb{I}\}, \qquad (40)$$

where $\rho_{AB} = \sum_i p_i \Psi_i' \otimes \Psi_i$ is given by the particular SE problem. There it is also shown that any separable decomposition of the optimal operator $\Lambda_{AB} = \sum_x M_x \otimes \phi_x$ corresponds to the optimal strategy (M_x, ϕ_x).

Now, consider the sequence of optimization problems:

$$\begin{aligned} F^N &\equiv \max\{\mathrm{tr}(\rho_{AB}\Lambda_{AB}) : \Lambda_{AB} \in S^N, \Lambda_A = \mathbb{I}\}, \\ F_p^N &\equiv \max\{\mathrm{tr}(\rho_{AB}\Lambda_{AB}) : \Lambda_{AB} \in S_p^N, \Lambda_A = \mathbb{I}\}, \end{aligned} \qquad (41)$$

From (3), it is immediate that $F^1 \geq F^2 \geq F^3 \geq ... \geq F$, with $\lim_{N\to\infty} F^N = F$. An analogous property holds for the bounds F_p^N. Note that these maximizations are SDPs and therefore can be easily computed.

Unfortunately, given limited computational (and specially memory) resources, it is only possible to compute these bounds up to some index N. In spite of the asymptotic convergence of the sequence, F^N or F_p^N could very well be far away from the actual solution of the problem. Is there any way to estimate the error of the truncation?

Take $\Lambda_{AB} \in S^N(S_p^N)$ to be the operator that maximizes equation (41). Theorem 2 (3) then implies that $\tilde{\Lambda}_{AB}$, as defined by equation (7) ((10)), corresponds to a feasible state estimation strategy, since it is separable and $\tilde{\Lambda}_A = \Lambda_A = \mathbb{I}$. Moreover, we can use the separable decomposition of $\tilde{\Lambda}_{AB}$ that appears in Section VI to express it as a measure-and-prepare protocol (M_x, ϕ_x).

The fidelities $\tilde{F}^N$ or $\tilde{F}_p^N$ associated to these strategies, although non trivial, will not be optimal in general, but they should provide a lower bound for F. From (40), it is easy to see that

$$\begin{aligned} \tilde{F}^N &= \frac{N}{N+d}F^N + \frac{1}{N+d}, \\ \tilde{F}_p^N &= \left(1 - \frac{dg_N}{2(d-1)}\right)F_p^N + \frac{g_N}{2(d-1)}. \end{aligned} \qquad (42)$$

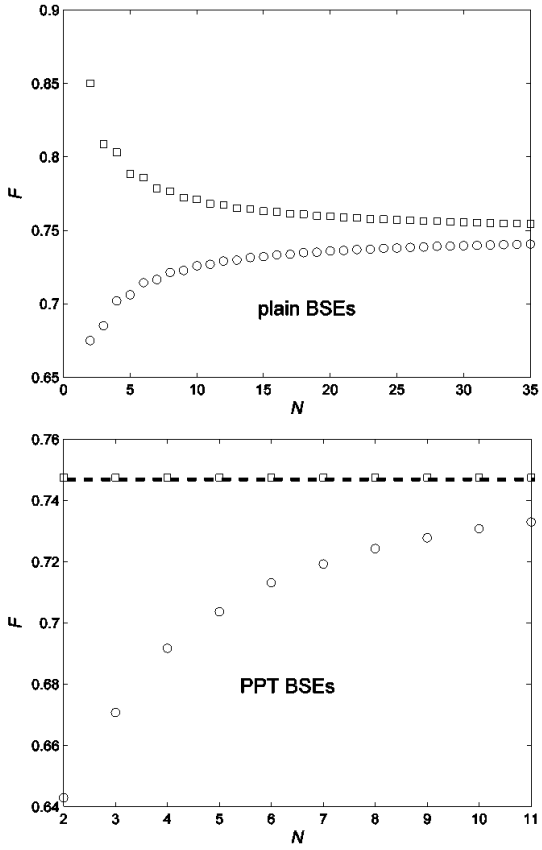

FIG. 1: Upper (squares) and lower (circles) bounds for the maximum fidelity F as a function of N. The dashed line indicates the value of the exact solution, attained exactly by the PPT upper bounds on F from $N = 2$ and onwards. The minimum difference between the upper and lower bounds is of the order of 10^{-2} in both plots.

Notice that both lower bounds asymptotically converge to F. That is, from the solutions of the semidefinite programs (41) it is possible to obtain a sequence of state estimation strategies that converges to the optimal measure-and-prepare scheme.

To have a grasp on the efficiency of the method, consider the following state estimation problem: suppose we have a device that outputs two copies of one of the 4 qubit states $\{|\Psi_k\rangle\}_{k=1}^4 \equiv \{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$ with equal probabilities. Our task is to estimate the state produced by the device. However, due to the environmental noise, once we are ready to measure the copies, those have degraded into $\rho_k \equiv \Omega^{(\epsilon)}(|\Psi_k\rangle\langle\Psi_k|) = (1-\epsilon)|\Psi_k\rangle\langle\Psi_k| + \epsilon\mathbb{I}/2$. The results for $\epsilon = 0.3$ are shown in Figure 1, for both the PPT and non PPT case and different values of N.

We used the MATLAB package *YALMIP* [47] in combination with *SeDuMi* [15] to perform the numerical calculations. Note that the curve corresponding to the upper bounds is constant, i.e., $F^N = F^M = F^*$, for all M, N. This suggested that F^* could be equal to F, the solution of the problem, although we did not observe any rank loop in the matrices output by the computer. We thus had to *force* the rank loop to occur. Using rank minimization heuristics [48] we checked for the existence of low rank PPT BSEs of Λ_{AB} such that $\mathrm{tr}(\Lambda_{AB}\rho_{AB}) \geq F^* - \delta$. Taking $\delta = 10^{-4}$, the computer returned a matrix with a rank loop, therefore proving the optimality of F^* up to this precision.

We performed a similar analysis for $d = 3$, this time considering the problem where a degraded copy of one of the states

$$|\psi_{ij}\rangle = \cos\left(\frac{j\pi}{6}\right)|0\rangle + \sin\left(\frac{j\pi}{6}\right)\cos\left(\frac{i\pi}{6}\right)|1\rangle + \\ + \sin\left(\frac{j\pi}{6}\right)\sin\left(\frac{i\pi}{6}\right)|2\rangle, \tag{43}$$

(where i and j run from 0 to 5) is sent to us with probability 1/36 through a depolarizing channel $\rho \to \Omega^{(0.2)}(\rho)$. In this case we were also able to force a rank loop in the PPT BSEs, so we again knew the optimal solution. Figure 2 illustrates our numerical results.

Note that, in both cases, the lower bounds on the solution behave very similarly as the upper bounds given by the DPS criterion, as long as we are considering the non PPT case. In the PPT case, however, our bounds prove to be terrible, since the second available upper bound obtained through the DPS criterion already seems to attain the optimal solution. We will discuss briefly this topic in Section IX.

The main features of the practical performance of the DPS criterion have already been illustrated above. Therefore, in the following two problems we will just stick to analytical results.

B. Maximal output purity of quantum channels

Let ω be a quantum channel. The *maximal output purity* [29] ν of ω is defined as

$$\nu = \max_{\rho} \|\omega(\rho)\|_\infty, \tag{44}$$

where the maximization is to be performed over all normalized quantum states ρ.

At first sight this quantity may seem extremely non linear. We will show that, actually, (44) can be reformulated as a linear optimization over the set of separable states.

Denote by Ω_{AB} the Choi operator corresponding to ω, i.e., $\omega(\rho) = \mathrm{tr}_A(\Omega_{AB} \cdot \mathbb{I}_A \otimes \rho)$. It follows that

$$\nu = \max_{\rho} \|\omega(\rho)\|_\infty = \max_{\rho,\sigma} \mathrm{tr}(\Omega_{AB} \cdot \sigma \otimes \rho), \tag{45}$$

with $\sigma, \rho \geq 0, \mathrm{tr}(\rho) = \mathrm{tr}(\sigma) = 1$.

Or, equivalently,

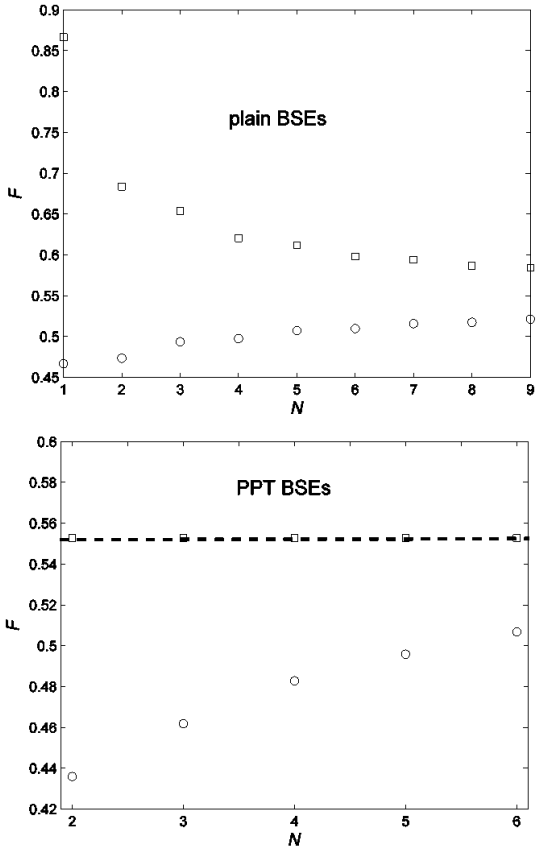

FIG. 2: Upper (squares) and lower (circles) bounds for the maximum fidelity F as a function of N in dimension 3. This time, the minimum difference between our lower bounds and the exact solution is around 0.03 (and it is attained in the non PPT case).

$$\nu = \max\{\mathrm{tr}(\Omega_{AB}\Lambda_{AB}) : \Lambda_{AB} \in \bar{S}\}. \tag{46}$$

As in the state estimation case, it is possible to define decreasing sequences $(\nu^N)_N, (\nu_p^N)_N$ of upper bounds on ν that converge asymptotically to the optimal output purity of the channel. Using Theorems 2 and 3, together with the fact that $\mathrm{tr}_B(\Omega_{AB}) = \mathbb{I}_A$, we have that there exist sequences $(\tilde{\nu}^N)_N, (\tilde{\nu}_p^N)_N$ of lower bounds on ν given by

$$\begin{aligned} \tilde{\nu}^N &= \frac{N}{N+d}\nu^N + \frac{1}{N+d}, \\ \tilde{\nu}_p^N &= \left(1 - \frac{dg_N}{2(d-1)}\right)\nu_p^N + \frac{g_N}{2(d-1)}. \end{aligned} \tag{47}$$

C. Geometric entanglement of tripartite pure states

Let $|\Psi\rangle_{ABC}$ be a pure tripartite state. A popular entanglement measure for this kind of systems is the so called *geometric entanglement* [30, 31] (that some mathematicians may recognize as the square of the *ϵ-norm* [32]), defined as

$$E = \max_{\phi_A,\phi_B,\phi_C} |\langle\phi_A|\langle\phi_B|\langle\phi_C||\Psi_{ABC}\rangle|^2. \tag{48}$$

Notice, though, that, if we fix ϕ_A and ϕ_B, the state ϕ_C maximizing the overlap will have to be proportional to $\langle\phi_A|\langle\phi_B||\Psi_{ABC}\rangle$. This overlap will be therefore equal to

$$\begin{aligned} &\mathrm{tr}_C(\langle\phi_A|\langle\phi_B||\Psi_{ABC}\rangle\langle\Psi_{ABC}||\phi_A\rangle|\phi_B\rangle = \\ &= \mathrm{tr}(\rho_{AB}|\phi_A\rangle\langle\phi_A| \otimes |\phi_B\rangle\langle\phi_B|), \end{aligned} \tag{49}$$

where $\rho_{AB} = \mathrm{tr}_C(|\Psi_{ABC}\rangle\langle\Psi_{ABC}|)$. It follows that E can also be reformulated as a linear optimization over S, i.e.,

$$E = \max\{\mathrm{tr}(\Lambda_{AB}\rho_{AB}) : \Lambda_{AB} \in \bar{S}\}. \tag{50}$$

As before, converging and decreasing sequences $(E^N)_N, (E_p^N)_N$ of upper bounds on E can be derived via the DPS criterion, and Theorems 2, 3 allow us to obtain complementary increasing sequences of lower bounds $(\tilde{E}^N)_N, (\tilde{E}_p^N)_N$, given by

$$\begin{aligned} \tilde{E}^N &= \tfrac{N}{N+d}E^N + \tfrac{1}{N+d}\lambda_A, \\ \tilde{E}_p^N &= \left(1 - \tfrac{dg_N}{2(d-1)}\right)E_p^N + \tfrac{g_N}{2(d-1)}\lambda_A. \end{aligned} \tag{51}$$

Here λ_A denotes the smallest eigenvalue of ρ_A.

VI. PROOF OF THEOREMS 2, 3

The purpose of this section is to derive Theorems 2, 3. But first, a few words on notation.

Given a unitary operator U, by $|U\rangle$ we will denote the state $U|0\rangle$. Also, for any permutation $\pi \in P_N$, $V_\pi \in B(\mathcal{H}^{\otimes N})$ will represent the corresponding permutation operator. V alone must be understood as the SWAP operator acting over a bipartite system $\mathcal{H}^{\otimes 2}$, i.e.,

$$V = \sum_{i,j=0}^{d} |i\rangle|j\rangle\langle j|\langle i|. \tag{52}$$

To finish, $\mathcal{H}_{\mathrm{sym}}^N$ will denote the symmetric subspace of $\mathcal{H}^{\otimes N}$ (the dimension of $\mathcal{H}$ will be clear from the context).

We will now proceed to proof Theorems 2, 3.

The basic idea for both proofs is to notice that the original problem of finding a separable state [62] $\tilde{\Lambda}_{AB}$ very

close to Λ_{AB} from its BSE Λ_{AB^N} can be viewed as a *probabilistic state estimation problem* [33].

Consider the following protocol, in which Alice plays a passive part:

1. A copy of Λ_{AB^N} is distributed to two parties, Alice and Bob.
2. Bob performs performs an incomplete measurement over $\mathcal{H}_B^{\otimes N}$, described by the POVM $\{M_x \geq 0\}_x$, with $\sum_x M_x \leq \mathbb{I}$. As a result, he obtains either an outcome x or a *FAIL message*, indicating that his measurement has failed to produce an outcome.
3. If Bob receives a FAIL message, then he makes it public. Otherwise, he prepares a state $\sigma_x \in B(\mathcal{H}_B)$, and both Alice and Bob would output the state $\frac{\mathrm{tr}_{B^N}(M_x \Lambda_{AB^N}) \otimes \sigma_x}{p_x}$ with probability $p_x = \mathrm{tr}_{B^N}(M_x \Lambda_{B^N})$.

The state Alice and Bob will produce conditioned on a non FAIL message will be then given by

$$\tilde{\Lambda}_{AB} = \sum_x \frac{\mathrm{tr}_{B^N}(M_x \Lambda_{AB^N}) \otimes \sigma_x}{\sum_y p_y}, \tag{53}$$

and is, therefore, a separable state. Moreover, since any entanglement breaking map can be decomposed as a measurement followed by the preparation of a state, this is the most general linear map we can apply over $\mathcal{H}_B^{\otimes N}$ in order to return a separable state $\tilde{\Lambda}_{AB}$.

But how to find a measure-and-prepare strategy for Bob such that $\tilde{\Lambda}_{AB}$ is close to Λ_{AB}? A possible scheme could be that Bob *pretended* that his subsystems are N identical copies of an unknown pure state, performed tomography over each of these subsystems independently and then prepared a state consistent with the average values he would measure. This strategy should give good results in the particular case where Λ_{AB^N} can be approximated by a state of the form

$$\int p(U) dU \rho_U \otimes |U\rangle\langle U|^{\otimes N}. \tag{54}$$

However, supposing that the state had the form above, an even better strategy would be to allow Bob to perform *collective* measurements over his subsystems and then prepare the most convenient state.

In conclusion, Bob should apply a POVM that allows him to efficiently identify the state $U|0\rangle\langle 0|U^\dagger$ out of N copies of it. Because in principle Bob has no a priori knowledge of $p(U)dU$, it is reasonable that he assumes that $p(U)dU = dU$, the Haar measure.

In this particular case, the best state estimation strategy and the best probabilistic state estimation strategy coincide [33]. This implies that Bob should apply the POVM $\{|U\rangle\langle U|^{\otimes N} dU\}_U$ and prepare the state $|U\rangle\langle U|$ whenever he gets the result U. Therefore,

$$\tilde{\Lambda}_{AB} = \frac{\int dU \mathrm{tr}_{B^N}\left(\mathbb{I}_A \otimes |U\rangle\langle U|^{\otimes N+1} \Lambda_{AB^N} \otimes \mathbb{I}_B\right)}{\int dU \mathrm{tr}(|U\rangle\langle U|^{\otimes N} \rho_{B^N})}. \tag{55}$$

To evaluate these integrals it is enough to notice that

1. For any operator C,

$$\int dU U^{\otimes N} C (U^\dagger)^{\otimes N} = \sum_{\pi \in P_N} c_\pi V_\pi, \tag{56}$$

for some coefficients c_π. In particular,

$$\int dU |U\rangle\langle U|^{\otimes N} = \frac{(d-1)! N! P^N_{\mathrm{sym}}}{(N+d-1)!} = \\ = \frac{(d-1)! \sum_{\pi \in P_N} V_\pi}{(N+d-1)!}. \tag{57}$$

2. Due to the fact that Λ_{AB^N} acts over $\mathcal{H}_A \otimes \mathcal{H}^N_{\mathrm{sym}}$, for any $\pi \in P_{N+1}$,

$$\mathrm{tr}_{B^N}\{(\Lambda_{AB^N} \otimes \mathbb{I}_B)\mathbb{I}_A \otimes V_\pi\} = \\ = \begin{cases} \Lambda_A \otimes \mathbb{I}_B, \text{ if } \pi(N+1) = N+1; \\ \Lambda_{AB}, \text{ otherwise.} \end{cases} \tag{58}$$

Finally, we arrive at the expression

$$\tilde{\Lambda}_{AB} = \frac{N}{N+d} \Lambda_{AB} + \frac{1}{N+d} \Lambda_A \otimes \mathbb{I}_B. \tag{59}$$

We have just proven Theorem 2.

The next step is to extend the previous ideas to account for the PPT condition, and a possible way is to modify the previous bipartite protocol to give Bob the ability to transpose part of his state before proceeding with any measure-and-prepare scheme. Suppose then that Bob partially transposes a partition B', corresponding to half of Bob's systems in Λ_{AB^N} (we will take N even for simplicity). Following the previous arguments, Bob could pretend that he and Alice are sharing a state $\Lambda_{AB}^{T_{B'}}$ very similar to

$$\int p(U) dU \rho_U \otimes (|U\rangle\langle U| \otimes |U^*\rangle\langle U^*|)^{\otimes N/2}. \tag{60}$$

The benefits of this apparently useless step become evident when we take into account the well established fact that it is easier to estimate a state from a copy and its complex conjugate than from two identical copies [33, 34]. In the case of $N = 2$, the optimal POVM has the form $\{U \otimes U^* |\varphi\rangle\langle\varphi| (U \otimes U^*)^\dagger dU\}$, where $|\varphi\rangle$ is a linear combination of $|00\rangle$ and $|\Psi^+\rangle = \sum_i |ii\rangle$, the (non normalized)

maximally entangled state. The optimal strategy for general N is not known, but we suggest the measurement

$$\phi_U dU \equiv (U \otimes U^*)^{\otimes N/2} |\phi\rangle\langle\phi| (U^\dagger \otimes (U^*)^\dagger)^{\otimes N/2} dU, \quad (61)$$

followed by the preparation of $|U\rangle\langle U|$. Here $|\phi\rangle$ is an arbitrary linear combination of the states [63] $|\phi_n\rangle \equiv |00\rangle^{\otimes n} |\Psi^+\rangle^{N/2-n}$, i.e.,

$$|\phi\rangle = \sum_{n=0}^{N/2} c_n |\phi_n\rangle. \quad (62)$$

Of course, applying the POVM ϕ_U over $\Lambda_{AB}^{T_{B'}}$ is equivalent to apply the (non positive!) map associated to $U^{\otimes N}|\phi\rangle\langle\phi|^{T_{B'}}(U^\dagger)^{\otimes N/2}$ over our state Λ_{AB^N}. That way, we can use the same tricks employed in the computation of (55).

A fast way to perform these calculations is to notice that, for $m > n$,

$$|\phi_n\rangle\langle\phi_m|^{T_{B'}} = |00\rangle\langle 00|^{\otimes n} \otimes (\mathbb{I} \otimes |0\rangle\langle 0|)^{\otimes m-n} \otimes V^{\otimes N/2-m}. \quad (63)$$

Therefore, there exists a pair of permutations $\pi, \pi' \in P_N$ such that

$$V_\pi |\phi_n\rangle\langle\phi_m|^{T_{B'}} V_{\pi'}^\dagger = |0\rangle\langle 0|^{\otimes m+n} \otimes \mathbb{I}^{\otimes N-m-n}. \quad (64)$$

But $\mathbb{I}_A \otimes V_\pi^\dagger \Lambda_{AB^N} = \Lambda_{AB^N} \mathbb{I}_A \otimes V_\pi = \Lambda_{AB^N}$, so

$$\begin{aligned} &\mathrm{tr}_{B^N}(\Lambda_{AB^N} \mathbb{I}_A \otimes U^{\otimes N} |\phi_n\rangle\langle\phi_m|^{T_{B'}} (U^\dagger)^{\otimes N}) = \\ &= \mathrm{tr}_{B^N}(\Lambda_{AB^N} \mathbb{I}_A \otimes |U\rangle\langle U|^{\otimes m+n} \otimes \mathbb{I}^{\otimes N-m-n}). \end{aligned} \quad (65)$$

In the end, we have that

$$\tilde{\Lambda}_{AB} = \left(1 - d \frac{\vec{c}^\dagger \tilde{A} \vec{c}}{\vec{c}^\dagger \tilde{B} \vec{c}}\right) \Lambda_{AB} + \frac{\vec{c}^\dagger \tilde{A} \vec{c}}{\vec{c}^\dagger \tilde{B} \vec{c}} \Lambda_A \otimes \mathbb{I}_B, \quad (66)$$

where $\tilde{A}$ and $\tilde{B}$ are square matrices given by

$$\begin{aligned} \tilde{B}_{nm} = \tfrac{(n+m)!}{(n+m+d-1)!}, \tilde{A}_{nm} = \tfrac{(n+m)!}{(n+m+d)!}, \\ n, m = 0, 1, ..., N/2. \end{aligned} \quad (67)$$

In case of odd N, we would make Bob partially transpose $(N-1)/2$ parts of his state and then use the following (incomplete) POVM:

$$U^{\otimes N} |\phi\rangle\langle\phi|^{T_{B'}} \otimes |0\rangle\langle 0| (U^\dagger)^{\otimes N} dU. \quad (68)$$

After the appropriate computations, we again arrive at expression (3), but the form of $\tilde{A}$ and $\tilde{B}$ changes to:

$$\begin{aligned} \tilde{B}_{nm} = \tfrac{(n+m+1)!}{(n+m+d)!}, \tilde{A}_{nm} = \tfrac{(n+m+1)!}{(n+m+d+1)!}, \\ n, m = 0, 1, ..., (N-1)/2. \end{aligned} \quad (69)$$

Obviously, in order to guarantee that Λ_{AB} is close to $\tilde{\Lambda}_{AB}$, it is in our interest to minimize the quantity

$$f_N(\vec{c}) \equiv \frac{\vec{c}^\dagger \tilde{A} \vec{c}}{\vec{c}^\dagger \tilde{B} \vec{c}} \quad (70)$$

over all possible vectors $\vec{c}$. Details on how to calculate the minimum of (70), together with the expression of the optimal $\vec{c}$ can be found in Appendix A. The result is:

$$\min_{\vec{c}} f_N(\vec{c}) = \frac{1}{2(d-1)} g_N. \quad (71)$$

This concludes the proof of Theorem 3.

Notice that in both cases the given separable decomposition of the states $\tilde{\Lambda}_{AB}$ is continuous. Because of the presence of the Haar measure, however, via Design Theory it is possible to arrive at an approximate [35] or exact [36] finite separable decomposition for these operators.

VII. EXTENSIONS TO MULTISEPARABILITY

So far, we have only been considering separability in *bipartite* systems. In this section, we show that almost all the results we have derived can be easily extended to deal with separability in m-partite scenarios. More concretely, we will show how to generalize Theorems 2 and 3 to the multipartite case, since, as we have already seen, most of the other results are just corollaries of these two theorems.

In this case, we will be interested in sets S^N of states that derive from an N *locally* (PPT) Bose-symmetric extension[37].

Definition 8. *N locally Bose-symmetric extension Let $\Lambda_{123...} \in \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3 \otimes ...)$ be a non negative operator. We will say that $\Lambda_{12^N 3^N ...} \in \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2^{\otimes N} \otimes \mathcal{H}_3^{\otimes N} \otimes ...)$ is an N locally Bose symmetric extension of $\Lambda_{123...}$ iff:*

1. $\Lambda_{12^N 3^N ...} \geq 0$.
2. $tr_{2^{N-1} 3^{N-1} ...}(\Lambda_{12^N 3^N ...}) = \Lambda_{123...}$.
3. $\Lambda_{12^N 3^N ...}$ *is independently Bose symmetric in systems* $2, 3, 4....$

As before, in case such extension is PPT with respect to some partition, we will denote it as an *N PPT locally Bose-symmetric extension.*

How close is $\Lambda_{123...}$ to the set of separable states? Consider a triseparable system, for instance, and suppose that we have an N locally Bose-symmetric extension

$\Lambda_{AB^NC^N}$ for Λ_{ABC}. In order to estimate the distance of Λ_{ABC} to the set of triseparable states we could conceive a protocol where the state $\Lambda_{AB^NC^N}$ is distributed between Alice, Bob and Charlie. As before, Bob and Charlie could then independently apply probabilistic state estimation over their subsystems and prepare both a quantum state depending on their measurement outcomes.

From what we already have, the derivation of the final expression of the triseparable state $\tilde{\Lambda}_{ABC}$ is straightforward. Equation (7) describes the action of Bob's strategy over *any* bipartite state. Considering the partition $AC^N|B^N$, it follows that the resulting tripartite state after Bob performs state estimation will be:

$$\frac{N}{N+d_B}\Lambda_{ABC^N}+\frac{1}{N+d_B}\Lambda_{AC^N}\otimes\mathbb{I}_B. \qquad (72)$$

Now it is Charlie's turn. This time we will take the partition $AB|C^N$. The final result is that

$$\tilde{\Lambda}_{ABC}=\tfrac{N^2}{(N+d_B)(N+d_C)}\Lambda_{ABC}+\tfrac{N}{(N+d_B)(N+d_C)}\Lambda_{AB}\otimes\mathbb{I}_C+ \\ +\tfrac{N}{(N+d_B)(N+d_C)}\Lambda_{AC}\otimes\mathbb{I}_B+\tfrac{1}{(N+d_B)(N+d_C)}\Lambda_A\otimes\mathbb{I}_{BC} \quad (73)$$

is a triseparable state.

The generalization to more parties is immediate. Invoking again the definition of depolarizing channels (4), in m-partite separability the expression for $\tilde{\Lambda}_{1234...}$ would be

$$\tilde{\Lambda}_{1234...}=(\mathbb{I}_1\bigotimes_{i=2}^{m}\Omega_{(p_i)})(\Lambda_{1234...}), \qquad (74)$$

where

$$p_i=\frac{d_i}{N+d_i}. \qquad (75)$$

The corresponding expression for $\tilde{\Lambda}_{123...}$ when it arises from an N locally Bose-symmetric extension, PPT with respect to the partition $12^{\lceil N/2\rceil}3^{\lceil N/2\rceil}...|2^{\lfloor N/2\rfloor}3^{\lfloor N/2\rfloor}...$, is still (74), but this time

$$p_i=\frac{d_i}{2(d_i-1)}g_N^{(d_i)}. \qquad (76)$$

VIII. THE POWER OF PPT ALONE

The Peres-Horodecki criterion, aka the PPT (Positive Partial Transpose) criterion [2], is one of the most popular existent criteria for entanglement detection. It is simple, it provides a very good approximation to the set of separable states in small dimensional cases and it usually leads to analytical results when applied over families of quantum states. Actually, some entanglement measures, like the negativity [39] or the PPT entanglement robustness [1] are based on the PPT condition.

It is interesting, thus, to try to determine how good the PPT criterion is for entanglement detection *alone*, i.e., not in combination with Doherty et al.'s method. Here, through a very simple argument, we show what we believe is the first result in this direction after the seminal paper of the Horodeckis [58].

The main idea of our derivation stems from the fact that positivity under partial transposition is equivalent to separability in $\mathbb{C}^3\otimes\mathbb{C}^2$ systems [58]. Suppose, then, that we have a PPT state $\rho_{AB}\in B(\mathcal{H}_A\otimes\mathcal{H}_B)$, with $d_A\geq 3$, and $d_B\geq 2$, and consider the (non normalized) state $\tilde{\rho}_{AB}$ given by

$$\tilde{\rho}_{AB}\propto\int dUdW P_U^3\otimes P_W^2\rho_{AB}P_U^3\otimes P_W^2, \qquad (77)$$

where dU and dW denote the Haar measures corresponding to $S(d_A)$ and $SU(d_B)$, respectively, and

$$P_U^3\equiv U\sum_{k=0}^{2}|k\rangle\langle k|U^\dagger, P_W^2\equiv W\sum_{k=0}^{1}|k\rangle\langle k|W^\dagger. \qquad (78)$$

It follows that $\tilde{\rho}_{AB}$ is a convex combination of unnormalized states $\rho_{U,W}\equiv P_U^3\otimes P_W^2\rho_{AB}P_U^3\otimes P_W^2$, with $\rho_{U,W}\in B(\mathbb{C}^3\otimes\mathbb{C}^2)$. Notice, also, that each $\rho_{U,W}$ is PPT, since

$$\rho_{U,W}^{T_B}= (P_U^3\otimes P_W^2\rho_{AB}P_U^3\otimes P_W^2)^{T_B}= \\ = P_U^3\otimes P_{W^*}^2\rho_{AB}^{T_B}P_U^3\otimes P_{W^*}^2\geq 0. \qquad (79)$$

Since PPT equals separability in $\mathbb{C}^3\otimes\mathbb{C}^2$ systems, it follows that each $\rho_{U,W}$ is separable, and so is $\tilde{\rho}_{AB}$, since by construction it is a convex combination of these states.

It only rests to find an analytical expression for $\tilde{\rho}_{AB}$. Using the previous techniques it is straightforward to arrive at

Theorem 9. *Let $\rho_{AB}\in B(\mathcal{H}_A\otimes\mathcal{H}_B)$ be a PPT normalized quantum state, with $d_A\geq 3, d_B\geq 2$. Then, for*

$$p_A=\frac{d_A(d_A-3)}{d_A^2-1}, p_B=\frac{d_B(d_B-2)}{d_B^2-1}, \qquad (80)$$

the state $\Omega^{(p_A)}\otimes\Omega^{(p_B)}(\rho_{AB})$ is separable.

Note that, in the particular case $d_A=3, d_B=2$, $\tilde{\rho}_{AB}=\rho_{AB}$.

By simple application of the tools already developed, we end up with the following Corollary.

Corollary 10. *For any PPT state ρ_{AB}, with $d_A\geq 3, d_B\geq 2$,*

$$R_G(\rho_{AB})\leq\frac{1}{12}(d_A+1)(d_B+1)-1, \qquad (81)$$

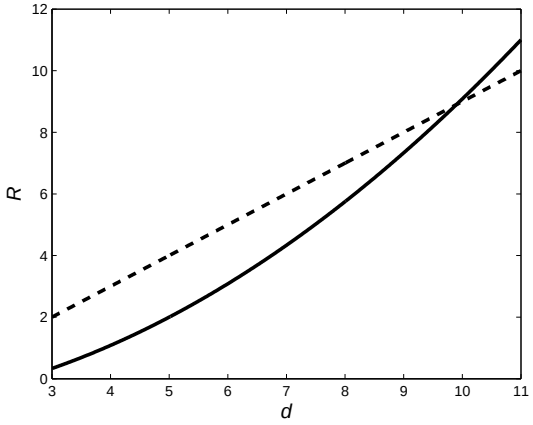

FIG. 3: Optimum bound on the global robustness of entanglement R for generic states (dashed line), as opposed to the upper bound for PPT states (solid line). In this plot, we assume that $d_A = d_B = d$. Note that the new bound becomes trivial as soon as $d > 9$.

and there exists a separable state σ such that

$$\|\rho_{AB} - \sigma\|_1 \leq 2 - \frac{24}{(d_A+1)(d_B+1)}. \qquad (82)$$

To get an idea on how good these bounds are, have a look at Figure 3. There the maximum possible *global* robustness of entanglement of a $\mathbb{C}^d \times \mathbb{C}^d$ state is compared with our upper bound for PPT states. We see that, although our upper bound becomes useless for $d > 9$, it is very powerful in the small dimensional case. For instance, for $\mathbb{C}^3 \times \mathbb{C}^3$ systems, the bound is equal to $1/3$ as opposed to 2. This means that we would have to apply the non PPT version of the DPS method up to $N = 6$ in order to characterize likewise the set of separable states.

IX. CONCLUSION

In this paper, we have studied the efficiency of the DPS criterion for entanglement detection. First, we showed that it is enough to subject the DPS states to some local noise in order to deprive them from their entanglement properties. It turned out that, while the minimal amount of noise necessary to turn an arbitrary state in $\bar{S}^N$ into a separable state decreases as $O(1/N)$, the corresponding amount of noise needed to disentangle states in $\bar{S}_p^N$ decreases at least as $O(1/N^2)$. We used these expressions to estimate the time complexity of both methods when applied to solve the Weak Membership Problem of Separability, and concluded that the PPT condition is worth imposing provided that the *optimal* bounds on the speed of convergence of the method based on plain BSEs scale as $O(d/N)$, as our own bounds suggest. We therefore hope to have shed some light on the question of how much the DPS criterion owes its strength to the PPT condition.

We also derived bounds on the error we incur when we substitute the set of separable operators by S^N or S_p^N in linear optimization problems, like the state estimation problem, the problem of determining the maximal output purity of an arbitrary quantum channel and the computation of the geometric entanglement. We performed numerical calculations of the first of these problems to test the accuracy of our analytical bounds. In order to compare our uncertainty with the actual solution of the problem, we developed a new technique that allows to prove in some cases the optimality of the DPS relaxations. We observed that, although the bounds for the non PPT case seem to be very accurate, the bounds for the PPT case are too big when compared with reality.

This disagreement between theory and practice may be explained in part by the fact that our bounds do not take into account the dimensionality of Alice's system, a crucial fact when dealing with the PPT constraint [58]. For all we know, our PPT bounds could be exact in the limit $d_A \to \infty$. Our intuition, nevertheless, is that better bounds could be found by applying linear maps over the initial state ρ_{AB} in order to obtain a separable state $\tilde{\rho}_{AB}$, as we did, but whose separable decomposition would be given by a *non linear map*, unlike in our examples. Actually, we already used that approach in Section VIII to bound the entanglement of PPT states. That kind of schemes, together with state estimation considerations, may allow in the future to obtain such better bounds.

Acknowledgements

The authors thank Animesh Datta and Fernando G. S. L. Brandão for useful discussions. This work is part of the EPSRC QIP-IRC and is supported by EPSRC grant EP/C546237/1, the Royal Society, the EU Integrated Project QAP and an Institute for Mathematical Sciences postdoc fellowship.

APPENDIX A: MINIMIZATION OF (70)

Take N even. Then it can be checked that

$$\tilde{A}_{mn} = \int_0^1 x^{m+n} \cdot \frac{(1-x)^{d-1}}{(d-1)!} dx,$$
$$\tilde{B}_{mn} = \int_0^1 x^{m+n} \cdot \frac{(1-x)^{d-2}}{(d-2)!} dx. \qquad (A1)$$

Combining this relation with (70), it follows that

$$f(\vec{c}) = \frac{1}{d-1} \frac{\int_0^1 |\sum_{n=0}^{N/2} c_n x^n|^2 (1-x)(1-x)^{d-2} dx}{\int_0^1 |\sum_{n=0}^{N/2} c_n x^n|^2 (1-x)^{d-2} dx}. \qquad (A2)$$

That way, we can see the minimization of $f(\vec{c})$ as a minimization over the set of all polynomials $Q_{N/2}(x) = \sum c_n x^n$ of degree $N/2$. Making the change of coordinates $y = 2x - 1$ we find that the above minimization is equivalent to

$$\min_{Q_{N/2}} \frac{1}{2(d-1)} \frac{\int_{-1}^{1} |Q_{N/2}(y)|^2 (1-y)^{d-1} dy}{\int_{-1}^{1} |Q_{N/2}(y)|^2 (1-y)^{d-2} dy}, \tag{A3}$$

where $Q_{N/2}(y)$ is an arbitrary polynomial of order $N/2$. This problem can be solved by means of the *Jacobi polynomials*.

The Jacobi polynomials $P_n^{(\alpha,\beta)}(y)$ are a complete set of functions orthogonal upon integration in the interval $[-1, 1]$ under the weight $(1+y)^\beta (1-y)^\alpha$ [21]. Now, define the *normalized Jacobi polynomials* $p_n(y)$ as

$$p_n(y) \equiv \frac{P_n^{(d-2,0)}(y)}{\|P_n^{(d-2,0)}\|}, \tag{A4}$$

with

$$\|P_n^{(d-2,0)}\| = \sqrt{\int_{-1}^{1} |P_n^{(d-2,0)}(y)|^2 (1-y)^{d-2} dy}. \tag{A5}$$

It is clear that we can express any $Q_{N/2}(y)$ as a linear combination of normalized Jacobi polynomials of order less or equal than $N/2$. That is,

$$Q_{N/2}(y) = \sum_{n=0}^{N/2} e_n p_n(y), \tag{A6}$$

for some coefficients e_n. Because of the orthogonality of the p_n's, when we input this expression in the integral of the denominator, we end up with

$$\int_{-1}^{1} |Q_{N/2}(y)|^2 (1-y)^{d-2} dy = \sum_n |e_n|^2. \tag{A7}$$

To calculate the integral on the numerator, we can make use of the recurrence relation

$$(1-y)p_n(y) = \alpha_n p_n(y) + \beta_n p_{n+1}(y) + \gamma_n p_{n-1}(y), \tag{A8}$$

that holds for some coefficients $\alpha_n, \beta_n, \gamma_n$, with $\gamma_0 = 0$ and $\gamma_{n+1} = \beta_n$ [21]. Invoking again the orthogonality of the Jacobi polynomials, we have that

$$\min_{\vec{c}} f(\vec{c}) = \min_{|\vec{e}|^2=1} \frac{1}{2(d-1)} \vec{e}^\dagger \tilde{C} \vec{e}, \tag{A9}$$

where $\tilde{C}$ is an $(N/2+1)\times(N/2+1)$ tridiagonal hermitian matrix given by

$$\tilde{C}_{m,n} = \begin{cases} \alpha_n, \text{ if } m = n, \\ \beta_n, \text{ if } m = n+1, \\ \gamma_n, \text{ if } m = n-1, \\ 0 \text{ elsewhere.} \end{cases} \tag{A10}$$

Now we will proceed to diagonalize $\tilde{C}$.

Let λ be an eigenvalue of $\tilde{C}$. This means that there exists a vector $\{v_i\}_{i=0}^{N/2+1}$ such that

$$(\alpha_n - \lambda)v_n + \beta_n v_{n+1} + \gamma_n v_{n-1} = 0, \tag{A11}$$

with $v_{N/2+1} = 0$.

Choose a real number y_0 and try the ansatz $v_n = p_n(y_0)$. From (A8), it is clear that v_n will satisfy (A11), provided that

$$\begin{aligned} \lambda &= 1 - y_0, \\ p_{N/2+1}(y_0) &= 0. \end{aligned} \tag{A12}$$

That is, any root of the polynomial $p_{N/2+1}(y)$ corresponds to an eigenvalue of $\tilde{C}$.

But $p_{N/2+1}(y)$ has $N/2+1$ *simple* roots [21], so all the eigenvalues of $\tilde{C}$ are obtained using this strategy. It follows that

$$\min_{\vec{c}} f_N(\vec{c}) = \frac{1}{2(d-1)} \min\{1 - x : P_{N/2+1}^{(d-2,0)}(x) = 0\}. \tag{A13}$$

Let us remark that this is not the first time the zeros of the Jacobi polynomials naturally appear in state estimation problems [57].

The expression for the case of odd N can be derived in an analogous way taking into account that, this time,

$$\begin{aligned} \tilde{A}_{mn} &= \int_0^1 x^{m+n} \cdot \frac{x(1-x)^{d-1}}{(d-1)!} dx, \\ \tilde{B}_{mn} &= \int_0^1 x^{m+n} \cdot \frac{x(1-x)^{d-2}}{(d-2)!} dx. \end{aligned} \tag{A14}$$

APPENDIX B: OPTIMALITY CRITERION (RANK LOOPS)

For some problems involving linear optimizations over the set S, it may happen (see [56]) that a particular relaxation of the problem F^N turns out to coincide with F. In this appendix we will show how this optimality can sometimes be detected.

We will take inspiration from optimality detection in other hierarchies of semidefinite programs that appear in

scientific literature. Consider the hierarchy of semidefinite programs used in [44] for the calculation of the maximal violation of linear Bell inequalities. There the optimality of a relaxation is detected when the rank of the matrix generated by the computer is equal to that of some of its submatrices. Remarkably, we can find similar results in the hierarchies of semidefinite programs defined by Henrion and Lasserre to minimize real polynomials in a bounded region of $\mathbb{R}^n$ [45].

The corresponding result in this scenario is the following:

Lemma 11. *Let Λ_{AB^N} be a BSE of Λ_{AB}, PPT with respect to the partition $AB^K|B^{N-K}$. If*

$$rank(\Lambda_{AB^N}) \leq \max\{rank(\Lambda_{AB^K}), rank(\Lambda_{B^{N-K}})\} \quad \text{(B1)}$$

then Λ_{AB} is a separable operator.

Following [44], we will say that Λ_{AB^N} presents a *rank loop* when it fulfills condition (B1).

The proof of Lemma 11 follows trivially from an old result by Horodecky et al. [46]:

Theorem 12. *Let ρ_{AB} be a PPT bipartite quantum state. If*

$$rank(\rho_{AB}) \leq rank(\rho_A), \quad \text{(B2)}$$

then ρ_{AB} is a separable state.

See [46] for a proof.

The possibility of finding a rank loop in practice in cases where the optimization over the set S_p^N coincides with the optimization over S should not be surprising. Note that any (finite dimensional) separable state Λ_{AB} can be expressed as a finite convex combination of product states, i.e.,

$$\Lambda_{AB} = \sum_{i=1}^{K} p_i \rho_i \otimes |\psi_i\rangle\langle\psi_i|, \text{ with } p_i > 0, \forall i, \quad \text{(B3)}$$

with $|\psi_i\rangle\langle\psi_i| \neq |\psi_j\rangle\langle\psi_j|$, for $i \neq j$. Now, consider the PPT Bose symmetric extension of Λ_{AB} given by

$$\Lambda_{AB^N} = \sum_{i=1}^{K} p_i \rho_i \otimes |\psi_i\rangle\langle\psi_i|^{\otimes N}, \quad \text{(B4)}$$

Clearly, as N tends to infinity, the vectors $\{|\psi_i\rangle^{\otimes N}\}_i$ become orthogonal. It follows that $K^* \equiv \lim_{N\to\infty} \text{rank}(\Lambda_{AB^N})$ exists and is equal to $\sum_i \text{rank}(\rho_i)$. Being the rank a natural number, this implies that there is an M such that, for any $N > M$, $\text{rank}(\Lambda_{AB^M}) = \text{rank}(\Lambda_{AB^N}) = K^*$. That is, for any finite dimensional separable state there exists a PPT Bose symmetric extension with a rank loop.

Of course, the fact that for any separable state ρ_{AB} there exists a PPT BSE with a rank loop does not mean that our computer is going to return such an extension. Note, though, that, if at the same time we set our computer to the task of finding PPT BSEs of ρ_{AB} we also demand a rank minimization of these matrices (i.e., we look for PPT BSEs with minimal rank), at some point we will find a rank loop.

Unfortunately, rank minimization of positive semidefinite matrices with linear constraints is in general an NP-hard problem [14, 59]. There are, however, heuristics [60] that have proven to be very efficient for solving small-scale problems (that is, for small d).

[1] M. B. Plenio and S. Virmani, Quant. Inf. Comp. **7**, 1 (2007).
[2] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996)
[3] D. Bruss, J. Math. Phys. **43**, 4237 (2002)
[4] B. Terhal J. Theor. Comp. Sci. **287**, 313 (2002)
[5] A. Sen, U. Sen, M. Lewenstein, and A. Sanpera, *e-print arXiv:quant-ph/*0508032.
[6] L.M. Ioannou, Quant. Inf. Comp., **7**, 335 (2007)
[7] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *e-print arXiv:quant-ph/*0702225.
[8] O. Gühne, G. Tóth, *e-print arXiv:*0811.2803.
[9] A. C. Doherty, P. A. Parrilo, F. M. Spedalieri, Phys. Rev. Lett., **88**, 187904 (2002).
[10] A. C. Doherty, P. A. Parrilo, F. M. Spedalieri, Phys. Rev. A **69**, 022308 (2004).
[11] G. A. Raggio and R. F. Werner, Helvetica Physica Acta, **62**, 980 (1989).
[12] R. F. Werner, Lett. Math. Phys. **17**, 359 (1989).
[13] C.M. Caves, C.A. Fuchs, and R. Schack, J. Math. Phys. **43**, 4537 (2002).
[14] L. Vandenberghe and S. Boyd, SIAM Review **38**, 49 (1996).
[15] J.F. Sturm, *SeDuMi, a MATLAB toolbox for optimization over symmetric cones*, http://sedumi.mcmaster.ca.
[16] J. Eisert, P. Hyllus, O. Guehne, M. Curty, Phys. Rev. A **70**, 062317 (2004).
[17] F. G. S. L. Brandão, R. O. Vianna, Phys. Rev. Lett. **93**, 220503 (2004).
[18] F. Hulpke, D. Bruss, J. Phys. A: Math. Gen. **38**, 5573 (2005).
[19] R. König and R. Renner, J. Math. Phys. **46**, 122102 (2005)
[20] M. Christandl, R. König, G. Mitshison, and R. Renner, *e-print arXiv:quant-ph/*0602130.
[21] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, New York: Dover Publications (1972).
[22] G. Vidal and R. Tarrach, Phys. Rev. A **59**, 141 (1999).
[23] G. Vidal, R.F. Werner, Phys. Rev. A **65**, 032314 (2002).
[24] M. Hayashi, D. Markham, M. Murao, M. Owari, S. Vir-

mani, Phys. Rev. Lett. **96**, 040501 (2006).
[25] D. Cavalcanti, Phys. Rev. A **73**, 044302 (2006).
[26] M. B. Plenio, F. G. S. L. Brandão, Nature, Physics **4**, 873-877 (2008).
[27] M. Horodecki and P. Horodecki, Phys. Rev. A **59**, 4206 (1999).
[28] N.J. Cerf, C. Adami and R.M. Gingrich, Phys. Rev. A **60**, 898 (1999).
[29] G.G. Amosov, A.S. Holevo, and R.F. Werner, Probl. Inf. Transm. **36**, 305 (2001).
[30] A. Shimony, Ann. N. Y. Acad. Sci. **755**, 675 (1995).
[31] T. Wei and P. Goldbart, Phys. Rev. A. **68**, 042307 (2003).
[32] see e.g. A. Defant and K. Floret, *Tensor Norms and Operator Ideals*, North-Holland (1993).
[33] J. Fiurasek, New J. Phys. **8**, 192 (2006).
[34] N. Gisin and S. Popescu, Phys. Rev. Lett. **83**, 432 (1999).
[35] A. Ambainis and J. Emerson, Twenty-Second Annual IEEE Conference on Computational Complexity (CCC07), pp. 129140 (2007).
[36] A. Hayashi, T. Hashimoto, M. Horibe, Phys. Rev. A, **72**, 032325 (2006).
[37] A. C. Doherty, P. A. Parrilo, F. M. Spedalieri, Phys. Rev. A, **71**, 032333 (2005).
[38] S. Beigi, P. W. Shor, *e-print arXiv:*0902.1806.
[39] M.B. Plenio, Phys. Rev. Lett. **95**, 090503 (2005), J. Eisert, PhD Thesis, Potsdam (2002).
[40] L. Gurvits, *Proceedings of the thirth fifth ACM symposium on Theory of computing*, 10, New York, ACM Press (2003).
[41] Sevag Gharibian, *e-print arXiv:*0810.4507.
[42] L. M. Ioannou, B. C. Travaglione, D. C. Cheung, and A. K. Ekert, Phys. Rev. A, **70**, 060303(R) (2004).
[43] L. M. Ioannou and B. C. Travaglione, Phys. Rev. A, **73**, 052314 (2006).
[44] M. Navascués, S. Pironio and A. Acín, New J. Phys. 10, **7**, 073013 (2008).
[45] D. Henrion and J. B. Lasserre, IEEE Trans. Aut. Contr. **51**, 192 (2006).
[46] P. Horodecki, M. Lewenstein, G. Vidal and I. Cirac, Phys. Rev. A **62**, 032310 (2000).
[47] J. Löfberg, *YALMIP : A Toolbox for Modeling and Optimization in MATLAB*, http://control.ee.ethz.ch/~joloef/yalmip.php.
[48] R. Orsi, *LMIRank: software for rank constrained LMI problems*, http://users.rsise.anu.edu.au/~robert/lmirank/.
[49] C.W. Helstrom, *Quantum Detection and Estimation Theory*, Academic Press, New York, (1976).
[50] A.S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, North-Holland, Amsterdam, (1982).
[51] K. Hammerer, M.M. Wolf, E.S. Polzik and J.I. Cirac, Phys. Rev. Lett. **94**, 150503 (2005)
[52] A. Serafini, O.C.O. Dahlsten and M.B. Plenio, Phys. Rev. Lett. **98**, 170501 (2007)
[53] G. Adesso and G. Chiribella, Phys. Rev. Lett. **100**, 170503 (2008)
[54] J. Calsamiglia, M. Aspachs, R. Munoz-Tapia and E. Bagan, *e-print arXiv:*0807.5126.
[55] M. Owari, M.B. Plenio, E.S. Polzik, A. Serafini and M.M. Wolf, New J. Phys., **10**, 113014 (2008).
[56] M. Navascués, Phys. Rev. Lett. **100**, 070503 (2008).
[57] P. Rapcan, J. Calsamiglia, R. Muñoz-Tapia, E. Bagan, V. Buzek, *e-print arXiv:*0708.1086.
[58] M. Horodecki, P. Horodecki and R. Horodecki, Physics Letters A, **223**,1 (1996).
[59] J. David, *Algorimth analysis for robust controllers*, PhD thesis, Kat. Univ. Leuven, ESAT, 3001, Leuven, Belgium (1994).
[60] M. Fazel, PhD Thesis, Stanford University (2002).
[61] To understand why, write $\tilde{\rho}$ as a convex combination of product states, i.e., $\tilde{\rho} = \sum p_i |u_i\rangle\langle u_i| \otimes |v_i\rangle\langle v_i|$. Then, $\tilde{\rho}_A \otimes \mathbb{I} - \tilde{\rho} = \sum p_i |u_i\rangle\langle u_i| \otimes (\mathbb{I} - |v_i\rangle\langle v_i|)$. That is, $\tilde{\rho}_A \otimes \mathbb{I} - \tilde{\rho} = \rho_A \otimes \mathbb{I} - \tilde{\rho}$ is a separable operator.
[62] Along this proof, we will consider the operator Λ_{AB} to be a quantum state rather than a quantum operator, i.e., to be normalized. It is easy to see that, if (7) and (10) are separable states when Λ_{AB} is normalized, the very same expressions have to lead to separable operators with $\mathrm{tr}(\tilde{\Lambda}_{AB}) = \mathrm{tr}(\tilde{\Lambda}_{AB})$ if Λ_{AB} is not normalized.
[63] Note that a further symmetrization of these states over the particles in B' and B
B' would be more intuitive, but irrelevant, since the support of the state $\Lambda_{AB^N}^{T_{B'}}$ is in $\mathcal{H}_A \otimes \mathcal{H}_{\mathrm{sym}}^{N/2} \otimes \mathcal{H}_{\mathrm{sym}}^{N/2}$. That is, such symmetrization is automatically performed when we apply this POVM over $\Lambda_{AB^N}^{T_{B'}}$.